TECHNICAL NOTE

Published in Conservation Genetics Resources, DOI 10.1007/s12686-016-0532-0

Development of SNP and microsatellite markers for goldsinny wrasse (*Ctenolabrus rupestris*) from ddRAD sequencing data

Authors: Jansson E.*, Taggart J.B.[#], Wehner, S. [#], Dahle G.*, Quintela M.*, Mortensen S.*, Kvamme B.O.* & Glover, K. A.*[&]

[*]*Institute of Marine Research, 5817 Bergen, Norway*

[&] *Sea Lice Research Centre, Department of Biology, University of Bergen, Norway*

[#]*Institute of Aquaculture, School of Natural Sciences, University of Stirling, Stirling FK9 4LA, United Kingdom*

Corresponding author: Eeva Jansson, eeva.jansson@imr.no

Abstract

Wrasse (*Labridae*) species have been used as parasite cleaners in Atlantic salmon farming since the 1980s. However, their use has recently escalated, with millions now being introduced into salmon cages each year. Most fish are of wild origin, their exploitation potentially impacting native populations. Genetic information is urgently required to inform management decisions. We identified 174 microsatellite and 149 SNP markers from ddRAD sequence data. From these, 17 and 48 microsatellite and SNP markers respectively were validated by genotyping 150 goldsinny wrasse collected from five locations along the Norwegian and Swedish coasts. Two to 30 alleles were identified at the microsatellite loci, while gene diversity (H_e) ranged 0.101–0.907. All SNP loci were biallelic, with averaged H_e per locus ranging between 0.063 and 0.495.

Keywords: SNP, microsatellite, goldsinny wrasse, *Ctenolabrus rupestris*, cleaner fish

The goldsinny wrasse, *Ctenolabrus rupestris* (Linnaeus 1758) is one of six Labridae species inhabiting Scandinavian waters, and one of four used in commercial salmonid farms as cleaner fish (Skiftesvik et al. 2015). Large-scale mixed fisheries for cleaner fish currently operate in several countries. In Norway alone, over 21 million wrasses were caught in 2014 (~11.7 million being goldsinnies; Norwegian Directorate of Fisheries 2015, www.fiskeridir.no/). Heavy fishing pressure combined with long-distance translocations raises concern about the effect that these practices may have on wild populations that are both potentially overfished, and receive human-mediated gene-flow. Therefore, genetic markers are urgently needed to conduct genetic studies.

Goldsinny genomic DNA was extracted from fin tissues using the Qiagen DNeasy Blood & Tissue Kit. For SNP discovery a 'standard' ddRAD library was constructed using five DNA samples from each of four Norwegian populations; methodology described elsewhere (Manousaki et al. 2016). For microsatellite discovery a second ddRAD library was prepared using one fish from each of four populations, using more frequent cutting restriction enzymes. A microsatellite enrichment step (Techen et al. 2010) was incorporated using the following oligonucleotide baits $(GACA)_6$, $(GATA)_6$, $(GGAT)_6$, $(AGC)_8$, $(GGA)_8$, $(GAA)_8$, $(AAT)_8$. Both libraries were sequenced as part of an Illumina MiSeq run (v2 chemistry, 160 base paired end reads). Stacks software (v1.27; Catchen 2013) was used to identify SNPs

from paired-end reads (de novo assembly; key parameters m=6, M=2, n=1). Targeted microsatellites with ≥ 8 repeats were identified with spreadsheet searches.

Stacks analysis identified 1371 RAD loci containing one or two SNPs in at least 17 of the 20 samples. Allele frequencies per population were computed and a subset of potentially informative SNPs (149) selected (Supplementary material 1); i.e. having a minimum minor allele frequency difference among the four populations of 0.3 and possessing suitable flanking sequence for PCR assay. Similarly a list of 174 microsatellite sequences was produced (Supplementary material 2), ranked by repeat number and PCR potential. From these lists 52 SNP and 25 microsatellite sequences were randomly selected for marker testing and validation. Those markers showing promise from an initial analysis were thereafter screened in an additional 150 samples collected from five sites (n = 28-33) along the Swedish and Norwegian coast.

PCR primers for microsatellites were designed using Primer3 software (v. 2.3.4; Untergrasser et al. 2012) implemented in Geneious v 9.0.4 (Biomatters). Microsatellites were amplified in four multiplexes (Table 1), each multiplex (10 μl) comprising 50 ng DNA template, 1 X Buffer, 2 mM $MgCl_2$, 0.1 mM dNTPs, 0.05-2 μM each primer (Table 1) and 0.05 U GoTaq polymerase. PCR profiles for the different multiplexes varied only by number of cycles. An initial 4 min denaturation at 94 °C, was followed by 24/25/22 cycles (groups 1-2/3/4, respectively) of 50 s at 94 °C, 50 s at 58 °C, and 80 s at 72 °C, and a final extension at 72 °C for 10 min. Forward primers were fluorescently labelled, amplicons being screened on an ABI Prism 377 Genetic Analyzer, and genotypes scored using GeneMapper v5 (Applied Biosystems). SNP locus primer design, amplification and genotype calling was based on the Sequenom MassARRAY iPLEX Platform, as described by Gabriel et al. (2009). Selected SNP loci were analyzed in two assay groups (Supplementary material 3). Seventeen of 25 microsatellite loci, and 48 of 52 SNP loci gave reliable polymorphic genotypes, and were analyzed further.

Micro-Checker v.2.2.3; (van Oosterhout et al. 2004) was used to check microsatellite loci for possible null alleles, large allele drop outs and scoring errors due to stuttering. 95% confidence intervals for scoring errors and null alleles were calculated with 1000 randomizations. Hardy-Weinberg equilibrium (HWE) and genotypic linkage disequilibrium (LD) between loci tests for all loci were performed with exact tests (Markov chain; 10000 dememorisations, 20 batches, 5000 iterations per batch) in Genepop (v. 4.3; Raymond & Rousset 1995, Rousset 2008). Number of alleles (*A*), expected and observed heterozygosities

99 (H_e, H_o), and inbreeding coefficients (F_{IS}), were calculated using GenAlEx (v. 6.5; Peakall
100 and Smouse 2006, 2012).

101 In the pooled dataset including all individuals, HWE tests showed significant deviation
102 from expectations (P<0.05) for four microsatellite (Table 1) and four SNP loci (Table 2).
103 However, consistent and statistically significant deviation from HWE was observed only for
104 the microsatellite locus Cru092 across all five study populations, and in three out of five
105 populations for the locus Cru081 (Supplementary material 4). These two loci were also
106 repeatedly suggested to harbour possible null alleles by the Micro-Checker analysis (Cru092
107 in all populations, and Cru081 in four out of five populations; data not shown). Nine SNP
108 locus pairs were known to be tightly linked (both SNPs in the same fragment being assayed;
109 Table 2). In addition to these, significant linkage (P<0.0001) was found between two
110 microsatellite marker pairs (Cru142/Cru012 and Cru016/Cru082). Microsatellite loci over all
111 samples had on average 9.494(±0.606) alleles (Table 1), and the mean H_e was 0.661(±0.028).
112 Averaged expected heterozygosity over all SNP markers and samples was 0.334(±0.009).

Table 1 Summary of 17 polymorphic microsatellite markers for goldsinny wrasse

Locus	Primers (5'-3')	Multiplex group	Primer conc. (μM)	Repeat motif	Product size (bp)	*N/A*	*He*	*Ho*	F_{IS}	HWEP
Cru007	F: TTGGTGTGAGGAGAAAGTGC-VIC R: ACTCCTGCCTGTCTGTGTAT	1	0.1	(GACA)n	115-211	150/23	0.888	0.913	-0.029	NS
Cru077	F: GAATCCTACCGGTCATCAGC-PET R: CTTAAAGCCCCGACGTAGAG	1	0.07	(GAA)n	131-137	150/2	0.427	0.454	-0.063	NS
Cru081	F: ATCCTCACCCTGAAGGAGAC-FAM R: CGTTTCCAGTTCCTACCCAG	1	0.1	(CTT)n	172-292	149/30	0.865	0.693	0.199	***
Cru142	F: TTTAAAAAGGGCACAGGGCT-PET R: AATCATCTCCATCAGCAGGC	1	0.2	(CCT)n	174-203	149/10	0.712	0.637	0.106	NS
Cru016	F: AAAAAGGAGCTGGACAGGAC-NED R: GATCAGGTGCTCTTGACCTG	2	0.15	(CTGT)n	177-277	149/24	0.907	0.832	0.083	**
Cru026	F: CTACAACCTGTCGCACCTAC-FAM R: TAGGTGAGGTGTGAGACAGG	2	0.1	(CTGT)n	153-165	150/4	0.205	0.23	-0.122	NS
Cru092	F: TGTGTTGTCAGTGTGTGGTT-PET R: GGCTCAAACAGAAGCTCCTT	2	0.1	(CTT)n	143-177	120/14	0.748	0.374	0.500	***
Cru099	F: TTGTGTTAGCTTGTGTGCCT-VIC R: TCCAACACCTCCCTCTCTTT	2	0.05	(GGA)n	83-115	150/12	0.799	0.815	-0.020	NS
Cru112	F: GAGTGTTTGGGATGGAGGTG-VIC R: ACCAGAACCCAAACACATCC	2	0.05	(GGA)n	189-201	150/5	0.101	0.106	-0.057	NS
Cru012	F: GATAAAGCCGAAGAGCCCTC-FAM R: CGCATATGGTCATCCGTTCT	3	0.1	(GACA)n	126-198	150/18	0.895	0.868	0.030	NS
Cru018	F: GGGGCAGTTTAAGCTAGCAA-NED R: GCGTCAGTGACACCTAACA	3	0.1	(CTGT)n	146-250	149/25	0.896	0.865	0.035	NS
Cru065	F: CTGGACTCATCGCAAAGACA-PET	3	0.2	(GAA)n	102-138	149/12	0.824	0.782	0.052	NS

	R: TGACCTCGCTGATGTCAGTA									
Cru037	F: AGCCAAGGAGACAAAATGGT-VIC	4	0.07	(GATA)n	135-193	146/18	0.802	0.843	-0.050	NS
	R: AGACTGATCCAAAACAGCTACC									
Cru082	F: CCGCTTCCTTCTTCCTCTTC-VIC	4	0.2	(CTT)n	232-313	148/18	0.727	0.695	0.043	**
	R: TAGAGAGCGGGAGAGAGAGA									
Cru110	F: CAGGCCCATTAGTGTCAGAC-PET	4	0.1	(GGA)n	180-201	119/6	0.640	0.608	0.049	NS
	R: TGCATTCTGTTGTCAGCTGT									
Cru132	F: TGTGAGGGGTTCATACAGGT-FAM	4	0.1	(GGA)n	163-172	150/3	0.243	0.248	-0.021	NS
	R: AACCAACTTACAGCCAGCTC									
Cru140	F: TCTCGCATAGAGGAGTCCAG-NED	4	0.1	(CCT)n	151-169	149/6	0.566	0.589	-0.038	NS
	R: CCCCTGCTGCACATTCATAT									

N/A number of samples amplified/number of alleles, *He* expected heterozygosity, *Ho* observed heterozygosity, F_{IS} inbreeding coefficient, HWEP result for H-W equilibrium test; NS = non-significant, * $P<0.05$, ** $P<0.01$, *** $P<0.001$

114
115

Table 2 Summary of 48 polymorphic SNP markers for goldsinny wrasse

Locus_SNP position	Primers (5'-3')	Allele	*N*	MAF	*He*	*Ho*	F_{IS}	HWEP
Locus10068_101	F:ACGTTGGATGTTATTTCAAATGCCGCCGCC	A/G	149	0.219	0.339	0.274	0.191	NS
	R:ACGTTGGATGAGAGGAAAACCCCAAAACCG							
Locus10440_106	F:ACGTTGGATGTCATGATCTCAACATCCCAC	C/T	148	0.291	0.397	0.435	-0.096	NS
	R:ACGTTGGATGGCCATATTGATACACTGATG							
Locus10844_70	F:ACGTTGGATGTTCCTTTGAGTGACTCTCCC	A/G	149	0.311	0.421	0.423	-0.005	NS
	R:ACGTTGGATGTTCTTTCTTCCACTCACAC							
Locus10844_86	F:ACGTTGGATGTTCTTTCTTCCACTCACAC	A/T	149	0.313	0.413	0.383	0.074	NS
	R:ACGTTGGATGTTCCTTTGAGTGACTCTCCC							
Locus10854_59	F:ACGTTGGATGGCAACCCCATTCACCTAATC	A/G	150	0.160	0.261	0.171	0.346	**
	R:ACGTTGGATGTGATATTTACTGGCTTTGC							
Locus10854_87	F:ACGTTGGATGTGATATTTACTGGCTTTGC	C/G	138	0.476	0.476	0.580	-0.219	***
	R:ACGTTGGATGGCAACCCCATTCACCTAATC							
Locus11260_41	F:ACGTTGGATGGACCAAAAAGGGTAACTGGG	C/G	149	0.352	0.455	0.474	-0.040	NS
	R:ACGTTGGATGCATGCTGTGAATGTTCCCTC							
Locus12300_71	F:ACGTTGGATGCGGAATATCGATAGTTTAA	A/G	150	0.479	0.488	0.558	-0.142	NS
	R:ACGTTGGATGTCGATACAGGGATACAAGTG							
Locus12300_108	F:ACGTTGGATGTCGATACAGGGATACAAGTG	A/T	149	0.248	0.367	0.374	-0.021	NS
	R:ACGTTGGATGCGGAATATCGATAGTTTAA							
Locus13412_49	F:ACGTTGGATGCTGTTCCAAGGCTTTTAATG	C/T	149	0.148	0.250	0.271	-0.084	NS
	R:ACGTTGGATGCCTTTCATTTATTATCTCC							
Locus13412_95	F:ACGTTGGATGCCTTTCATTTATTATCTCC	A/G	147	0.049	0.092	0.097	-0.060	NS
	R:ACGTTGGATGCTGTTCCAAGGCTTTTAATG							
Locus13542_71	F:ACGTTGGATGAAAAAACGTCCCTGGCAGAG	A/G	147	0.427	0.487	0.405	0.168	*
	R:ACGTTGGATGTGACAGCTAGTGATGTTACC							
Locus13594_103	F:ACGTTGGATGCCTCTCCCAGCTTCTCCTT	C/G	150	0.130	0.217	0.260	-0.196	NS

	R:ACGTTGGATGCGTCAGCATGAATCTGTTTG							
Locus13663_99	F:ACGTTGGATGGGGTTATCTACTTTGAAATG	C/T	150	0.227	0.348	0.335	0.038	NS
	R:ACGTTGGATGGGATCTTCTGTGTTGACTGG							
Locus137_47	F:ACGTTGGATGTTACATATGCCTACCTCCCC	A/G	150	0.273	0.395	0.398	-0.006	NS
	R:ACGTTGGATGTCAAACCACCGAGGAAGAAG							
Locus13732_76	F:ACGTTGGATGGGTCACAGACTGAACACAAC	A/G	150	0.133	0.229	0.214	0.067	NS
	R:ACGTTGGATGATCCCAACGATCAGAGAGTG							
Locus15371_74	F:ACGTTGGATGCAATAGAATGATTGGACTAGC	G/T	146	0.279	0.395	0.259	0.345	***
	R:ACGTTGGATGGAGGCTGGATCCCCTTTTTG							
Locus1870_72	F:ACGTTGGATGCTTCGGAGTACACGTGAGAA	G/T	148	0.388	0.474	0.402	0.153	NS
	R:ACGTTGGATGGAATTGTTTAGCTGGCATCC							
Locus1883_66	F:ACGTTGGATGTACCTAGAACCACGACTGAC	C/T	149	0.287	0.401	0.327	0.184	NS
	R:ACGTTGGATGTGCTAATGTGCTGTGGTCTC							
Locus2038_99	F:ACGTTGGATGACCTCTGGAGGCCTTTTAAC	C/G	150	0.302	0.419	0.352	0.161	NS
	R:ACGTTGGATGACTGTAATTACCTGCAATC							
Locus207_103	F:ACGTTGGATGAGTGAGTCTCTGCGTGTCTG	A/G	150	0.292	0.410	0.378	0.079	NS
	R:ACGTTGGATGGGAAAAAGGTAGGCTAATCTC							
Locus3090_60	F:ACGTTGGATGTTATCTGATTACCCTGACGG	A/T	150	0.246	0.368	0.345	0.062	NS
	R:ACGTTGGATGGTTGAAGAAGAACAGGCTCC							
Locus3299_58	F:ACGTTGGATGGTTATTTATCTGTTTTATCCC	A/G	148	0.454	0.482	0.404	0.162	NS
	R:ACGTTGGATGTTAAAATGTCAGAGGAAAC							
Locus3299_101	F:ACGTTGGATGGTTATTTATCTGTTTTATCCC	A/T	150	0.120	0.208	0.188	0.094	NS
	R:ACGTTGGATGTTAAAATGTCAGAGGAAAC							
Locus3594_98	F:ACGTTGGATGCAAAGACAGCACCTATAAA	A/C	149	0.483	0.495	0.499	-0.007	NS
	R:ACGTTGGATGTGATTAGAACCATTTTAAAC							
Locus3684_35	F:ACGTTGGATGGTCCGCCTGTCTTGTTAACT	A/G	150	0.053	0.100	0.106	-0.065	NS
	R:ACGTTGGATGGCAGGAGTGTGTGTTCATTTT							
Locus3684_63	F:ACGTTGGATGGTCCGCCTGTCTTGTTAACT	C/G	147	0.159	0.263	0.237	0.099	NS

	R:ACGTTGGATGGCAGGAGTGTGTGTTCATTTT							
Locus3836_83	F:ACGTTGGATGGCTCTGACTTAAAGTTCACG	A/G	150	0.146	0.245	0.266	-0.084	NS
	R:ACGTTGGATGCTGTTCATGTGTTCTACAGG							
Locus4072_41	F:ACGTTGGATGTTCAGGGCGACACAAACCTC	A/C	150	0.342	0.436	0.456	-0.045	NS
	R:ACGTTGGATGAAGTCGACCTTTCCACTTCC							
Locus4263_103	F:ACGTTGGATGTGAACACTGTCAGTCCACAC	A/T	149	0.397	0.461	0.424	0.082	NS
	R:ACGTTGGATGCCCCTGTCAACACCATGAGA							
Locus4688_92	F:ACGTTGGATGCTGCAGTTTGTCTAAAACTTC	G/T	149	0.344	0.446	0.483	-0.082	NS
	R:ACGTTGGATGGTTAGTTTAGCAGCTTAGCAC							
Locus5363_67	F:ACGTTGGATGTTTGTTTATGCTGGTGTGCC	C/T	149	0.082	0.148	0.138	0.065	NS
	R:ACGTTGGATGCCACATTGGATGTCCAACAG							
Locus5704_64	F:ACGTTGGATGTCTGAATGTCAATGCCCCTC	A/G	150	0.234	0.314	0.361	-0.151	NS
	R:ACGTTGGATGATTAAGGTCCCATGGGCTTC							
Locus605_109	F:ACGTTGGATGACAGTATGCATCACTGGCTC	G/T	150	0.447	0.479	0.477	0.004	NS
	R:ACGTTGGATGAATGTAGAAGGTAGACGTG							
Locus6318_97	F:ACGTTGGATGCAGAAACTTGGAGAAGCTCG	A/G	148	0.163	0.266	0.246	0.078	NS
	R:ACGTTGGATGGCGTCTCACTTGTATTGCTG							
Locus6440_49	F:ACGTTGGATGTCCTGGGCTTCCTCTTTCTC	A/T	142	0.166	0.274	0.290	-0.057	NS
	R:ACGTTGGATGCAGAGCAGCGGGAGGAGCA							
Locus6440_72	F:ACGTTGGATGCAGAGCAGCGGGAGGAGCA	A/C	148	0.128	0.217	0.201	0.073	NS
	R:ACGTTGGATGTCCTGGGCTTCCTCTTTCTC							
Locus6583_45	F:ACGTTGGATGGACAGTCCTCCTACAATCAG	G/T	144	0.069	0.128	0.125	0.020	NS
	R:ACGTTGGATGCTCATCTGAGGAGGAACATC							
Locus6883_68	F:ACGTTGGATGAAAGATGCCATGACAGTGCC	A/G	148	0.033	0.063	0.066	-0.051	NS
	R:ACGTTGGATGGAGGCTGGATCCCCTTTTTG							
Locus713_94	F:ACGTTGGATGGACATAAAAACTGTCCAACCC	A/T	148	0.419	0.472	0.393	0.167	NS
	R:ACGTTGGATGCCTACAACACAAAGTGTAACG							
Locus7167_86	F:ACGTTGGATGGTAATCGTTGCTGTGCTTTG	A/G	149	0.075	0.134	0.108	0.190	NS

	R:ACGTTGGATGATCATCGCTCAGCCGGGTTT							
Locus739_47	F:ACGTTGGATGTTACGAACCATTCTGTCCTC	G/T	150	0.258	0.379	0.368	0.027	NS
	R:ACGTTGGATGGGTTTGTTAGATGGTGACGC							
Locus8493_91	F:ACGTTGGATGTGATCTTTCAGCTCAGGGTC	A/T	149	0.262	0.381	0.375	0.017	NS
	R:ACGTTGGATGATGGGTGGGTAAAGCAGTTG							
Locus8916_62	F:ACGTTGGATGCCGCTTGTTCTATATGATA	C/T	150	0.333	0.438	0.399	0.090	NS
	R:ACGTTGGATGGCGTCTGATGGAGGAAGAAA							
Locus8916_102	F:ACGTTGGATGCCGCTTGTTCTATATGATA	C/T	149	0.312	0.408	0.382	0.065	NS
	R:ACGTTGGATGGCGTCTGATGGAGGAAGAAA							
Locus9226_60	F:ACGTTGGATGAGGTGTCTTCGTCCCTTTTG	A/C	149	0.432	0.489	0.412	0.158	NS
	R:ACGTTGGATGCTTTGTGACGTCTGCTCTTG							
Locus9375_79	F:ACGTTGGATGATTCTACAGGGTCAGGTTG	A/C	148	0.229	0.349	0.324	0.072	NS
	R:ACGTTGGATGGGAAGATGTTTGAAACTTG							
Locus9375_108	F:ACGTTGGATGGGAAGATGTTTGAAACTTG	A/C	148	0.091	0.165	0.183	-0.107	NS
	R:ACGTTGGATGATTCTACAGGGTCAGGTTG							

N number of samples, MAF minor allele frequency, *He* expected heterozygosity, *Ho* observed heterozygosity, F_{IS} inbreeding coefficient, HWEP result for H-W equilibrium test; NS = non-significant, * P<0.05, ** P<0.01, *** P<0.001

Acknowledgements: This study was financed by the Norwegian Department for Trade and Fisheries (NFD). The Swedish Cultural Foundation in Finland (Svenska kulturfonden) is acknowledged for personal grant to EJ. The authors also acknowledge the support of the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland). The authors thank Paulo Prodohl for supplying biotinylated oligos.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

Catchen J, Hohenlohe P, Bassham S, Amores A, Cresko W. (2013) Stacks: an analysis tool set for population genomics. Mol. Ecol. 22: 3124–3140

Gabriel S, Ziaugra L, Tabbaa D (2009) SNP Genotyping Using the Sequenom MassARRAY iPLEX Platform. Curr Protoc Hum Genet 2.12.1-2.12.18

Kearse M, Moir R, Wilson A et al. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12):1647–1649

Manousaki T, Tsakogiannis A, Taggart JB et al. (2016) Exploring a non-model teleost genome through RAD sequencing - Linkage mapping in Common Pandora, *Pagellus erythrinus* and comparative genomic analysis. G3: Genes Genomes Genetics 6:509-519

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Raymond R, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. J Hered 86(3):248–249

Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8(1):103–106

Skiftesvik AB, Durif CMF, Bjelland RM, Browman HI (2015) Distribution and habitat preferences of five species of wrasse (Family Labridae) in a Norwegian fjord. ICES J Mar Sci 73(3):890–899

Techen N, Arias RS, Glynn NC, Pan Z, Khan, IA, Scheffler BE. (2010) Optimized construction of microsatellite-enriched libraries. Mol Ecol Resour 10: 508–515

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3--new capabilities and interfaces. Nucleic Acids Res 40(15):e115

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538